DESIGNING PERFORMANCE OF MICROPROCESSOR SPEED

Computer Organization Assignment

ABSTRACT

Three generations of Alpha microprocessors have been designed using a proven custom design methodology. The performance of these microprocessors was optimized by focusing on high-frequency design. The Alpha instruction set architecture facilitates high clock speed, and the chip organization for each generation was carefully chosen to meet critical paths. Digital has developed six generations of CMOS technology optimized for high-frequency design. Complex circuit styles were used extensively to meet aggressive cycle time goals. CAD tools were developed internally to support these designs. This paper discusses some of the technologies that have enabled Alpha microprocessors to achieve high performance.

Group 3

Ahirwe Marie Claudine 217079261

Ahishakiye Anaclet 217049021

Habimana Eliane 217046029

Niyitegeka David 217060730

Tuyizere Anastase 217038859

Muhire Arsene 217244904

Niyigena Jean Paul 217021891

Mugabo Francois 217187412

Designing Microprocessor Speed

Abstract:

Three generations of Alpha microprocessors have been designed using a proven custom design methodology. The performance of these microprocessors was optimized by focusing on high-frequency design. The Alpha instruction set architecture facilitates high clock speed, and the chip organization for each generation was carefully chosen to meet critical paths. Digital has developed six generations of CMOS technology optimized for high-frequency design. Complex circuit styles were used extensively to meet aggressive cycle time goals. CAD tools were developed internally to support these designs. This paper discusses some of the technologies that have enabled Alpha microprocessors to achieve high performance.

Terms:

Generally, the microprocessor is an integrated circuit and it incorporates core function of a computer's central processing unit. The microprocessor is a programmable multipurpose silicon chip, register based, clock driven, it accepts input as a binary data and after the processing, it provides the output data as per the instructions stored in the memory.

III. TECHNOLOGY

Digital Semiconductor has developed six generations of

CMOS process technology, with a new technology for each

major microprocessor design. The microprocessor design occurs

in parallel with the development of the manufacturing

process. Therefore, close cooperation is required between

the process development and microprocessor design teams

to perform this concurrent design and ensure optimum chip

performance. Table II highlights the key features of the three

process technologies used to produce these three microprocessors.

The processes were optimized for high-frequency

microprocessor design. In particular, emphasis is placed on

low ’s and very short ’s which increase drive current at

the cost of higher leakage.

Close interaction between the circuit design team and the process development team also results in the following benefits.

1) Early process information and timely updates of technology parameters are provided to the design teams, allowing circuit design to start before the process is fully defined.

2) Early design work provides valuable feedback to the process team to ensure that target process performance is met.

3) Major process features such as number of interconnect layers, interconnect pitch, and device characteristics are managed in the context of the overall chip design.

4) The design of critical structures such as RAM arrays and data paths can be optimized through process and circuit design.

A. Definition of Design Rules

One of the key areas where close collaboration is required between design and process development teams is the definition of layout design rules. Aggressive design rules can result in increased circuit density, and can potentially improve overall chip performance. However, design rules that are too aggressive will complicate manufacturing, and may impact yield. On the other hand, slack design rules may result in increased die size, resulting in increased distances between critical structures. This increased distance results in higher capacitance, larger routing delays, and lower chip performance.

Often, the process team can be more aggressive if limits are placed not only on the minimum widths and spaces of structures, but also on the maximum widths and spaces. The 21264 implements metal fillers to limit the maximum spacing between adjacent lines. The fill metal is automatically placed in the design and tied to or . For large areas of fill metal, stress relief holes are automatically placed in the fill pattern. Metal fill may increase the capacitance of nearby signal lines, but they also result in improved interlayer dielectric uniformity. The improved uniformity allows the process to be targeted more aggressively. Fig. 5 shows filler polygons inserted in the gaps between widely spaced lines.

Accordingly, while the chipmakers have been busy learning how to fabricate chips of greater and greater density, the processor designers must come up with ever more elaborate techniques for feeding the monster. Among the techniques built into contemporary processors are the following:

• Branch prediction: The processor looks ahead in the instruction code fetched from memory and predicts which branches, or groups of instructions, are likely to be processed next. If the processor guesses right most of the time, it can prefetch the correct instructions and buffer them so that the processor is kept busy.The more sophisticated examples of this strategy predict not just the next branch but multiple branches ahead. Thus, branch prediction increases the amount of work available for the processor to execute.

• Data flow analysis: The processor analyzes which instructions are dependent on each other's results, or data, to create an optimized schedule of instructions.

In fact, instructions are scheduled to be executed when ready, independent of the original program order. This prevents unnecessary delay.

• Speculative execution: Using branch prediction and data flow analysis, some processors speculatively execute instructions ahead of their actual appearance in the program execution, holding the results in temporary locations. This enables the processor to keep its execution engines as busy as possible by executing instructions that are likely to be needed.

These and other sophisticated techniques are made necessary by the sheer power of the processor.They make it possible to exploit the raw speed of the processor. Nowhere is the problem created by such mismatches more critical than in the interface between processor and main memory. Consider the history depicted in Figure 2.10.While processor speed has grown rapidly, the speed with which data can be transferred between main memory and the processor has lagged badly.The interface between processor and main memory is the most crucial pathway in the entire computer because it is responsible for carrying a constant flow of program instructions and data between memory chips and the processor. If memory or the pathway fails to keep pace with the processor's insistent demands, the processor stalls in a wait state, and valuable processing time is lost.

There are a number of ways that a system architect can attack this problem, all of which are reflected in contemporary computer designs. Consider the following examples:

• Increase the number of bits that are retrieved at one time by making DRAMs "wider" rather than "deeper" and by using wide bus data paths.

• Change the DRAM interface to make it more efficient by including a cache7 or other buffering scheme on the DRAM chip.

• Reduce the frequency of memory access by incorporating increasingly complex and efficient cache structures between the processor and main memory.

This includes the incorporation of one or more caches on the processor chip as well as on an off-chip cache close to the processor chip.

• Increase the interconnect bandwidth between processors and memory by using higher-speed buses and by using a hierarchy of buses to buffer and structure data flow.

Another area of design focus is the handling of I/O devices. As computers become faster and more capable, more sophisticated applications are developed that support the use of peripherals with intensive I/O demands. Figure 2.11 gives some

Performance Balance While processor power has raced ahead at breakneck speed, other critical components of the computer have not kept up.The result is a need to look for performance balance:

Nowhere is the problem created by such mismatches more critical than in the interface between processor and main memory. Consider the history depicted in Figure 2.10.While processor speed has grown rapidly, the speed with which data can be transferred between main memory and the processor has lagged badly.The interface between processor and main memory is the most crucial pathway in the entire computer because it is responsible for carrying a constant flow of program instructions and data between memory chips and the processor. If memory or the pathway fails to keep pace with the processor's insistent demands, the processor stalls in a wait state, and valuable processing time is lost.2.2 / DESIGNING FOR PERFORMANCE 41

There are a number of ways that a system architect can attack this problem, all of which are reflected in contemporary computer designs. Consider the following examples:

• Increase the number of bits that are retrieved at one time by making DRAMs "wider" rather than "deeper" and by using wide bus data paths.

• Change the DRAM interface to make it more efficient by including a cache7 or other buffering scheme on the DRAM chip.

• Reduce the frequency of memory access by incorporating increasingly complex

and efficient cache structures between the processor and main memory.
This includes the incorporation of one or more caches on the processor chip as well as on an off-chip cache close to the processor chip.

• Increase the interconnect bandwidth between processors and memory by using higher-speed buses and by using a hierarchy of buses to buffer and structure data flow.

Another area of design focus is the handling of I/O devices. As computers become faster and more capable, more sophisticated applications are developed that support the use of peripherals with intensive I/O demands. Figure 2.11 gives some examples of typical peripheral devices in use on personal computers and workstations.These devices create tremendous data throughput demands.While the current generation of processors can handle the data pumped out by these devices, there remains the problem of getting that data moved between processor and peripheral.
Strategies here include caching and buffering schemes plus the use of higher-speed interconnection buses and more elaborate structures of buses. In addition, the use of multiple-processor configurations can aid in satisfying I/O demands.
The key in all this is balance. Designers constantly strive to balance the throughput and processing demands of the processor components, main memory, I/O devices, and the interconnection structures. This design must constantly be rethought to cope with two constantly evolving factors:

• The rate at which performance is changing in the various technology areas (processor, buses, memory, peripherals) differs greatly from one type of element to another.

• New applications and new peripheral devices constantly change the nature of the demand on the system in terms of typical instruction profile and the data access patterns.

Thus, computer design is a constantly evolving art form. This book attempts to present the fundamentals on which this art form is based and to present a survey of the current state of that art.

References:

1. High-Performance Microprocessor Design Paul E. Gronowski, Member, IEEE, William J. Bowhill, Member, IEEE, Ronald P. Preston, Member, IEEE, Michael K. Gowan, and Randy L. Allmon, Associate Member, IEEE
2. https://www.efxkits.co.uk/different-types-of-microprocessors-uses/